Journal of Number Theory **77**, 326–330 (1999)
Article ID jnth.1999.2389, available online at http://www.idealibrary.com on IDEAL®

On the Dimension of the Hilbert Cubes

Norbert Hegyvári*

Department of Mathematics, ELTE TFK, Eötvös University, Markó u. 29, H-1055 Budapest, Hungary
E-mail: NORB@ludens.elte.hu

Communicated by A. C. Woods

Received October 30, 1997

A sequence of positive integers with positive lower density contains a Hilbert (or combinatorial) cube size $c \log \log n$ up to n. We prove that $c \log \log n$ cannot be

In [1] D. Hilbert showed (using different terminology) that for any $k \geqslant 1$, if $\mathbf{N}$ is finitely colored than there exists in one color infinitely many translates of a k-cube. We call $H \subset \mathbf{N}$ a k-cube if there exist $a > 0$ and $x_1, x_2, ..., x_k$ such that

$$H = H(a, x_1, ..., x_k) = \left\{ a + \sum_{i=1}^{k} \varepsilon_i a_i : \varepsilon_i = 0 \text{ or } 1 \right\}.$$

This result was essentially the first Ramsey-type theorem. The density version of the Hilbert result is the following:

Theorem A: *Let $k \geqslant 1$ be a integer and assume $A \subset [1, n]$ and*

$$|A| \geqslant n^{1-2^{-k}}.$$

Then A contains a k-cube. (*see* [2], [3]).

This result implies the following

Corollary. *Let A be an infinite sequence of integers with*

$$d(A) = \lim_{n \to \infty} \inf A(n)/n > 0,$$

* Research partially supported by Hungarian National Foundation for Scientific Research, Grant T025617, and by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) Grant NSF-STC-91-19999.

0022-314X/99 $30.00

where $A(n)=\sum_{a_i \leqslant n} 1$. *Then there exists a* β *such that for every* n, $A \cap [1, n]$ *contains a* k*-cube where*

$$k > \beta \log \log n. \tag{$*$}$$

Remark. It is not too hard to detect that in the proof of Theorem A we can achieve that the k-cube would be non-degenerate, i.e., the vertices of the cube are distinct. Furthermore, it is easy to see by a counting method that in this case $(*)$, apart from the constant factor, cannot be improved.

In this note we allow the degenerate cube as well.

We prove the following theorem:

THEOREM. *There exists an infinite sequence* A *of positive integers with* $\underline{d}(A) > 0$ *and*

$$H(n) \leqslant c \sqrt{\log n \log \log n}$$

where $c = 4 \cdot (\log(4/3))^{-1/2}$.

The proof of the theorem is not constructive. A. Sárközy and the author investigated the dimension of the Hilbert cube for some conventional sequence as well. For instance, we prove that if Q is the sequence of squares then $H_Q < c' \sqrt[3]{\log n}$, and for the sequence of primes P we get $c_1 \cdot \log \log n < H_P(n) < c_2 \cdot \log n$ (see [6]). Finally, we mention a related question of P. Erdös solved by E. Straus (see [5]): Is it true that if A is an infinite sequence of integers and A has positive upper density then A contains an infinite Hilbert cube? The answer is negative.

For sequences A, B of integers let $A+B=\{a+b : a \in A, b \in B\}$. For the proof of the theorem we need some lemmas.

LEMMA 1. *Let* $B=\{b_1<b_2< \cdots <b_k\}$ *be a sequence of integers. Then*

$$\binom{k+1}{2}+1 \leqslant |\Sigma\{B\}| \leqslant 2^k,$$

where $\Sigma\{B\} = \{x \colon x=\sum_{i=1}^{k} \varepsilon_i b_i : \varepsilon_i = 0 \text{ or } 1, b_i \in B\}$.

The proof is easy or see [4].

LEMMA 2. *We have*

$$T=|\{A \colon A \subseteq [1, n];\ |A|=k \text{ and } |\Sigma(A)|<k^3\}| < n^{3 \log_2 k} \cdot 3^{k^2}.$$

Proof of Lemma 2. Let $U=\{A: A\subseteq[1,n]; |A|=k \text{ and } |\Sigma(A)|<k^3\}$. Let $R=[3\log_2 k]$ and assume $A=\{a_1<\cdots<a_k\}\in U$. An element a_j is called a *doubler* if

$$\Sigma\{a_1, a_2, ..., a_{j-1}\}\cap\{\{a_j\}+\Sigma\{a_1, a_2, ..., a_{j-1}\}\}=\{\emptyset\}. \tag{1}$$

Since

$$\Sigma\{a_1, a_2, ..., a_j\}=\{0, a_j\}+\Sigma\{a_1, a_2, ..., a_{j-1}\},$$

thus if a_j is doubler then

$$|\Sigma\{a_1, a_2, ..., a_j\}|=2\,|\Sigma\{a_1, a_2, ..., a_{j-1}\}|. \tag{2}$$

This yields that

$$|\Sigma\{a_1, a_2, ..., a_k\}|\geqslant 2^H \tag{3}$$

where H denotes the number of doublers in A.

H is at most R since in the opposite case $2^H\geqslant 2^{R+1}>2^{3\log_2 k}=k^3$, which by (3) contradicts the fact $A\in U$.

Now if a_j is not a doubler then we must have

$$a_j\in\{x-x' \mid x, x'\in\Sigma\{a_1, ..., a_{j-1}\}\},$$

which easily implies that we can write a_j in the form

$$a_j=\sum_{i\neq j}\delta_i a_i; \qquad \delta_i\in\{0, +1, -1\} \tag{4}$$

This yields that the number of non-doubler elements is at most 3^k.

Now we get an upper estimation for T.

We can select $\binom{k}{R}$ subscripts j where a_j is the doubler, the number of possible values of the doublers being at most n^R. Finally, the number of non-doublers is at most $(3^k)^{k-R}$. Thus we have

$$\begin{aligned} T &\leqslant \binom{k}{R}^R\cdot(3^k)^{k-R}\\ &\leqslant k^R\cdot n^R\cdot 3^{k^2-kR}\leqslant n^R\cdot 3^{k^2}\end{aligned}$$

using the inequality $k^R<3^{kR}$.

Lemma 2 is proved.

Now we turn to the proof of the Theorem.

Proof of the Theorem. Let X be a random sequence of integers with $\Pr(x\in X)=\frac{1}{16}$. Clearly with probability 1 we have $\underline{d}(X)>0$. Let H_n be the event $H_X(n) > c\sqrt{\log n \log\log n}$ where $c = 4 \cdot (\log 4/3)^{-1/2}$. We are going to show

$$\Pr(H_n) < \frac{1}{n^2} \tag{5}$$

We have

$$\begin{aligned}\Pr(H_n) &\leqslant \sum_{\substack{1 \leqslant a \leqslant n \\ 1 \leqslant x_1, ..., x_k \leqslant n}} \left(\frac{1}{16}\right)^{|\Sigma(x_1, ..., x_k)|} \\ &= \sum_{\substack{1 \leqslant a \leqslant n \\ 1 \leqslant x_1, ..., x_k \leqslant n \\ |\Sigma(x_1, ..., x_k)| < k^3}} \left(\frac{1}{16}\right)^{|\Sigma(x_1, ..., x_k)|} + \sum_{\substack{1 \leqslant a \leqslant n \\ 1 \leqslant x_1, ..., x_k \leqslant n, \\ |\Sigma(x_1, ..., x_k)| \geqslant k^3}} \left(\frac{1}{16}\right)^{|\Sigma(x_1, ..., x_k)|}.\end{aligned} \tag{6}$$

By Lemmas 1 and 2 we have

$$\sum_{\substack{1 \leqslant a \leqslant n \\ 1 \leqslant x_1, ..., x_k \leqslant n, \\ |\Sigma(x_1, ..., x_k)| < k^3}} \left(\frac{1}{16}\right)^{|\Sigma(x_1, ..., x_k)|} \leqslant \sum_{1 \leqslant a \leqslant n} n^{3\log_2 k} \cdot 3^{k^2} \cdot \left(\frac{1}{16}\right)^{k^2/2} = n \cdot n^{3\log_2 k} \cdot \left(\frac{3}{4}\right)^{k^2},$$

which is less than $1/2n^2$ if $k \geqslant 4 \cdot \log(4/3)^{-1/2}\sqrt{\log n \log\log n}$. Furthermore,

$$\sum_{\substack{1 \leqslant a \leqslant n \\ 1 \leqslant x_1, ..., x_k \leqslant n, \\ |\Sigma(x_1, ..., x_k)| \geqslant k^3}} \left(\frac{1}{16}\right)^{|\Sigma(x_1, ..., x_k)|} \leqslant n \cdot \binom{n}{k} \cdot \left(\frac{1}{16}\right)^{k^3} < \frac{1}{2n^2}$$

holds if $k > 3\sqrt{\log_2 n}$.

By (5) we have $\sum_{n=1}^{\infty} \Pr(H_n) < \infty$ so that by the Borel–Cantelli lemma with probability 1, at most a finite number of events H_n occur.

This completes the proof of the theorem.

Remarks. (1) We split the sum in (6) into two parts according the value of $|\sum(x_1, ..., x_k)|$. We mention here that for the sets $A_d = \{d, 2d, ..., kd\}$, $d = 1, 2, ..., [\frac{n}{k}]$, we have

$$\left|\sum(A_d)\right| = \binom{k+1}{2} \sim k^2.$$

So we have to count these sets in the first sum which yields that our method works only if $k > c\sqrt{\log n}$.

(2) There are many pairs A and A' for which $\sum(A) \cap \sum(A')$ contains a fixed "big" set which is why we use only the trivial inequality $\Pr(\bigcup A) \leqslant \sum \Pr(A)$.

In this section we are going to show that for a *random* sequence our bound, apart from the factor $\sqrt{\log \log n}$, is the best possible.

We prove

PROPOSITION. *Let A be a random sequence of positive integers with $\Pr(a \in A) = p$. Then with probability* 1, *we have*

$$H_A(n) > c_p \sqrt{\log n}.$$

Proof of the Proposition. Let $0 < p < 1$ be a fixed real number and let A be a random sequence of integers with $\Pr(a \in A) = p$ and let $k_n = \max_{a,k}\{k: a+1, ..., a+k \in A\}$. By a theorem of Erdös and Rényi [7], with probability 1 $k_n = c(p) \log n$. But let us observe that if $a, a+1, ..., a+k \in A$ then $H(a, 1, 2, ..., [\sqrt{2k} - 1]) \subset A$. (Indeed, $H(a, 1, 2, ..., [\sqrt{2k} - 1]) \supseteq [a, a+1, ..., k]$.) This yields that with probability 1 we have $H_A(n) > c_p \sqrt{\log n}$, where $c_p = \sqrt{2c(p)}$.

ACKNOWLEDGMENT

I thank Dr. I. Z. Ruzsa for his suggestions and remarks.

REFERENCES

1. D. Hilbert, Über die Irreducibilität rationaler Functionen mit ganzzäligen Koefficienten, *J. Reine Angew. Math.* **110** (1892), 104–109.
2. P. Erdös, A. Sárközy, and V. T. Sós, On a conjecture of Roth and related problems I, *Algorithms Combin.* **8**, 47–59.
3. R. L. Graham, B. Rothschild, and J. Spencer, "Ramsey Theory," Wiley Interscience, 1980.
4. H. Hegyvári, On representation problems in the additive number theory, *Acta Math. Hung.* **72** (1–2) (1996), 35–44.
5. N. Hindman, Ultrafilters and combinatorial number theory, *in* "Number Theory Carbonadle" (M. Nathanson, Ed.), Lecture Notes in Math., Vol. 751, pp. 119–184.
6. N. Hegyvári and A. Sárközy, On the size of a combinatorial cube of certain sets, to appear.
7. P. Erdös and A. Rényi, On a new law of large numbers, *J. Anal. Math.* **22** (1970), 103–111.